MULTIMEDIA UNIVERSITY

STUDENT ID NO

MULTIMEDIA UNIVERSITY

FINAL EXAMINATION

TRIMESTER 2, 2022/2023

BAP1014 – ANALYTICAL PROGRAMMING

(All sections / Groups)

3 JULY 2023

9.00A.M.- 11.00 A.M.
(2 Hours)

INSTRUCTIONS TO STUDENTS

1. This question paper consists of 7 pages (including the cover page) with 2 sections.
2. Answer **ALL** questions. All the distribution of the marks for each question is given.
3. Please write all your answers in the Answer Booklet and Graph Paper provided.

SECTION A (50 Marks)

Choose One Answer for Each Question Below (Each Question Caries 2 Marks)

1. Python Pandas is an open-source library?
   A. Yes
   B. No
2. Python Pandas was created by?
   A. Guido van Rossum
   B. Travis Oliphant
   C. Wes McKinney
   D. Brendan Eich
3. What is the purpose of Pandas in Python?
   A. Reading data
   B. Reshaping data
   C. Subsetting data
   D. All of the above
4. Pandas series is similar to a?
   A. To a row in a table
   B. To a column in table
   C. To a spreadsheet table
   D. All of the above
5. What is the default label for the third value in a panda's series?
   A. 0
   B. 1
   C. 2
   D. 3
6. Pandas DataFrame is similar to a?
   A. To a row in a table
   B. To a column in table
   C. To a spreadsheet table
   D. All of the above
7. Which of the following is the correct features of a DataFrame?
   A. Potentially columns are of different types
   B. Can Perform Arithmetic operations on rows and columns
   C. Labeled axes (rows and columns)
   D. All of the above
8. Which of the following is the correct component of a DataFrame?
   A. Series
   B. Rows
   C. Columns
   D. All of the above

... Continue

9. What is the data contained in a DataFrame called?
    A. Index
    B. Values
    C. Columns
    D. Label
10. What would be the result of the following code?
    fi = pd.DataFrame (Red, index = "blue", "blue", "white")
    A. The label of the red DataFrame will be changed to blue, blue, and white
    B. The values of the red DataFrame will be changed to blue, blue, and white
    C. The label of the Red DataFrame will be changed to blue, blue, and white
    D. The values of the Red DataFrame will be changed to blue, blue, and white
11. Which of the following can be a data in Pandas?
    A. A python dict
    B. An ndarray
    C. A scalar value
    D. All of the above
12. Which method allows for the bottom 15 rows of the DataFrame to be shown?
    A. head(15)
    B. sample (15)
    C. describe(15)
    D. tail(15)
13. Which method will allow for a random selection of rows to be shown?
    A. head()
    B. mean()
    C. sample()
    D. describe()
14. What does the loc method allows you to do in a DataFrame?
    A. Access a specified value
    B. Change a specified index
    C. Access the bottom row
    D. All of the above
15. Which Pandas methods can be used for Data Cleaning?
    A. dropna()
    B. describe()
    C. isnull()
    D. All of the above

... Continue

16. What is the purpose of Numpy in Python?
    A. For numerical calculation and scientific computing
    B. Multidimensional array
    C. Both A and B
    D. None of the above
17. Numpy inherently has slow operations of Array?
    A. True
    B. False
18. Numpy array can be?
    A. Flatten
    B. Sort
    C. Reshape
    D. All of the above
19. Numpy array cannot be more than 3-dimensional?
    A. True
    B. False
20. The function of Numpy concatenate method() is?
    A. Reshape array
    B. Create a copy of an array
    C. Join array along a new axis
    D. None of the above
21. What is the purpose of Matplotlib in Python?
    A. For data analysis
    B. For generating graphs
    C. For generating plots
    D. All of the above
22. Does Matplotlib not allow for 3D graphs?
    A. True
    B. False
23. Which of the following does not visualize data?
    A. Histogram
    B. Shapes
    C. Maps
    D. None of the above
24. Which library is built-on top of Matplotlib?
    A. Numpy
    B. SciPy
    C. Seaborn
    D. All of the above
25. Pandas, Numpy, and Matplotlib can be used together?
    A. True
    B. False

... Continue

SECTION B (50 Marks)

Answer the questions based on the program

Presented below is a Python program for **DATA ANALYSIS** called **SALES**. However, it is not a working program.

```
df = pd.read_csv('sales_data.csv')
df['Price'].fillna(df['Price'].mean(), inplace=True)
    z_scores = (df['Revenue'] - df['Revenue'].mean()) / df['Revenue'].std()
    df = df[(z_scores < 3)]
category = input("Enter the product category: ")
plt.show()
import matplotlib.pyplot as plt
if region:
df = df[df['Region'] == region]
df['Date'] = pd.to_datetime(df['Date'])
import pandas as pd
df['Quantity'].fillna(df['Quantity'].mean(), inplace=True)
import numpy as np
region = input("Enter the region: ")
df['Month'] = df['Date'].dt.month
monthly_revenue = df.groupby('Month')['Revenue'].sum()
    plt.plot(monthly_revenue.index, monthly_revenue.values)
plt.xlabel('Month')
    plt.ylabel('Total Revenue')
    plt.title(f"Monthly Revenue Trend for {category} in {region} Region")
if category:
df = df[df['Category'] == category]
    plt.xticks(range(1, 13))
plt.grid(True)
```

...Continue

1. Construct a workable program by **REARRANGING** the codes above in the **CORRECT ORDER AND FORMAT**.
   Note: Rewrite the whole program

(16 Marks)

2. Classify **STEPS** of the program and **INTERPRET** what each step does in a **TABLE FORMAT**.

(9 Marks)

Date	Region	Category	Quantity	Price	Revenue	Month
15/1/2023	West	Clothing	20	38.59509	772	1
18/1/2023	West	Clothing	48	95.79405	4598	1
23/1/2023	West	Clothing	87	22.63215	1969	1
21/2/2023	West	Clothing	7	80.41218	563	2
25/2/2023	West	Clothing	80	23.41289	1873	2
23/3/2023	West	Clothing	30	30.18233	905	3
31/3/2023	West	Clothing	50	12.74495	637	3
16/4/2023	West	Clothing	13	20.03516	260	4
24/4/2023	West	Clothing	99	36.40724	3604	4
29/4/2023	West	Clothing	60	83.55037	5013	4
1/5/2023	West	Clothing	96	26.57167	2551	5
7/5/2023	West	Clothing	72	31.73523	2285	5
9/5/2023	West	Clothing	11	14.96668	165	5
16/5/2023	West	Clothing	50	17.76029	888	5
1/6/2023	West	Clothing	5	51.62773	258	6
24/7/2023	West	Clothing	17	75.01795	1275	7
7/8/2023	West	Clothing	7	99.24254	695	8
22/8/2023	West	Clothing	56	73.80812	4133	8
31/8/2023	West	Clothing	19	97.98229	1862	8
9/9/2023	West	Clothing	29	98.2882	2850	9
3/10/2023	West	Clothing	85	93.35357	7935	10
21/10/2023	West	Clothing	69	11.75496	811	10
28/10/2023	West	Clothing	53	17.05276	904	10
15/11/2023	West	Clothing	27	21.79237	588	11
23/11/2023	West	Clothing	66	94.45449	6234	11
24/11/2023	West	Clothing	4	24.18782	97	11
27/11/2023	West	Clothing	7	21.65133	152	11
5/12/2023	West	Clothing	96	11.05174	1061	12
6/12/2023	West	Clothing	27	74.60069	2014	12
10/12/2023	West	Clothing	31	35.23733	1092	12
24/12/2023	West	Clothing	10	30.60586	306	12

... Continue

Shown above is the **DATA** for **SALES** with the Region == "West" and Category == "Clothing."

3. Interpret the **BUSINESS PROBLEM** of this dataset which could be solved by doing **DATA ANALYSIS** with evidence.

(4 Marks)

4. Construct a **GRAPH** based on the Sales program with the **GIVEN DATA** above.

(15 Marks)

5. Interpret how the **DATA ANALYSIS** done can **SOLVE THE BUSINESS PROBLEM** in question 3 section B with evidence.

(6 Marks)

End of Page